WEAVING A WEB: CONCEPT ACQUISITION AND INFERENTIAL ROLE

TECENDO UMA TEIA: AQUISIÇÃO DE CONCEITOS E PAPEL INFERENCIAL

John Sarnecki*

ABSTRACT – Theories of concepts and concept acquisition are mutually constraining. How we envisage concept acquisition depends both on what we take concepts to be and what skills we can employ to acquire them. I argue that Ned Block’s cognitivist approach to concept acquisition is not compatible with his vision of conceptual role semantics. If concepts are defined by their conceptual roles, then the acquisition of new concepts will change the conceptual roles of concepts employed in any form of hypothesis formation and confirmation learning. This breaks the evidentiary link between the concepts acquired and the evidence used to justify its subsequent applications. As a consequence, conceptual role semantics cannot avail itself of cognitivist approaches to concept acquisition. Despite this, they may nevertheless explain the apparent rational nature of much concept acquisition.


RESUMO – Teorias dos conceitos e da aquisição de conceitos são mutuamente vinculados. O modo como encaramos a aquisição de conceitos depende tanto do que tomamos por conceitos e das habilidades que empregamos para adquiri-los. Eu argumento que a abordagem cognitivista da aquisição de novos conceitos proposta por Ned Block não é compatível com a sua concepção semântica de papel conceitual. Se os conceitos são definidos pelos seus papéis conceituais, então a aquisição de novos conceitos mudará os papéis conceituais de conceitos empregados em qualquer forma de formação de hipótese e de aprendizagem de confirmação. Isso quebra a conexão que evidencia entre conceitos adquiridos e a evidência utilizada para justificar suas aplicações subsequentes. Consequentemente, a semântica de papel conceitual não pode beneficiar-se das abordagens cognitivistas para a aquisição de conceitos. A despeito disso, elas podem, entretanto, explicar a natureza aparentemente racional da aquisição de conceitos.


* University of Toledo, Ohio. <john.sarnecki@utoledo.edu>.

Veritas Porto Alegre v. 57 n. 3 set./dez. 2012 p. 138-162
Taking on the problem of meaning, or the nature of linguistic or mental representation, is no minor task. Modern theories of meaning have done much to chart out specific proposals about how individual tokens or symbols, instantiated in mind or language, may come to represent or in some sense “stand for” particular individuals or states of affairs.

Posing the question in modern naturalist garb, Blackburn writes, “[h]ow is it even possible for a world – a natural world of things in space and time, of flesh and blood – to contain some things that represent other things?” (BLACKBURN, 1984, 39) Blackburn’s question is at once about the metaphysical nature of representation and, in flesh and blood, the human capacity to employ it. We may say that in the first part the question is about what makes representation possible and in the second it is how we make it actual. If, as some might claim, representation is not a property of human minds, then the first question is considerably less interesting. But even granting this, we may suppose that how humans represent can only be part of the story about the nature of representation, since we are only one of many possible representation systems. However, if we are impressed, as Blackburn is, by the existence of meaningful expressions and thoughts in the physical world, then the goal of theory of a theory of meaning will be to explain how meaningful states are instantiated in this world, how they can be naturalized. A naturalized semantics is thus interested in more than how representation is possible. It is interested in how representation emerges from the physical world and how, ultimately, it is reducible to it.

In some ways, this analysis reverses the script. We are not simply spelling out the nature of meaning based on an analysis of the nature of representation. Instead, we can work backwards from the constraints on human representational capacities to determine the adequacy of theories of meaning more generally.

Hence, these accounts can be instructive. Most theories of mental representation are steeped in the language of concepts (see, for example, FODOR, 1981; PEACOCKE, 1992). A concept functions as a semantic atom, much like a word in a language. It is through the concatenation and combination of concepts (according to some pseudo-grammatical rules of combination) that we may describe particular states of affairs.

1 In recent years, this assumption has begun to encounter some significant scrutiny. While no current account does away with representation, recent work in dynamical systems theory and embodiment have threatened to undermine standard representational accounts of cognition. Even so, such views remain controversial both in terms of their critiques and the representational properties of their competing accounts. See, for example, Wheeler (2005) for one such account.
Concepts, on this model, underlie our capacity for systematic and productive thought. On these terms, then, a theory of concepts is at the heart of a theory of representation. In this paper, I want to look at one further constraint: concept acquisition.

A theory of concept acquisition mediates between two goals. On the one hand, a theory of concept acquisition must prove amenable to some account of what it is to have states with a specific representational content. As Christopher Peacocke notes, questions of concept possession are logically prior to questions of concept acquisition (see PEACOCKE, 1992, ix). Our theory of how we acquire a concept will depend on what we think concepts are. But a theory of concept acquisition must also negotiate a path between these constraints on the nature of concepts and the empirically determined constraints of a theory of instantiation. We cannot accept a constitutive theory of concepts that makes the possession of these concepts impossible. If your theory says that concepts are bowling balls, then, short of discovering bigger heads (or smaller bowling balls), we don’t have any concepts. These constraints are sharpened within a theory of concept acquisition, for not only must we offer a credible picture of the instantiation of a given model of concept possession we must also show that there is some possible progression of steps by which we come to instantiate these states. Naturalism commits us to the view that any theory of this progression must also be empirically credible. The bowling ball theory, for example, may be implausible based on the size of our heads; it is even more so based on the possible avenues into them.

Semantic Holism

Navigating between these competing demands has defined much of the debate about the nature of concepts. In this paper, I will examine what has been perhaps the dominant approach to the naturalization of mental representations: inferential role semantics. Proponents of this view have maintained that it is fundamentally the logical or causal relationships between mental representations that account for the semantic properties of mental states. Variants of this view can be traced as far as Wittgenstein’s claim that meaning is a function of how a symbol is used. However, there has been a broad proliferation of theories corresponding to different ways of defining the circumstances of use and the relations between mental representations within specific systems. These can range from Brandom’s (1994) profoundly normative and social account of these relations to psychological accounts of concepts in terms of the interrelation of theoretical terms in a child’s
emerging scientific understanding of the world (GOPNIK and MELTZOFF, 1997).

The focus of this discussion will be Ned Block’s (1986) justly famous paper, *An Advertisement for a Semantics for Psychology*. Block’s theory remains relevant in this discussion both because of its seminal nature – few inferential roles theories developed today have not been influenced by the main tenets of Block’s account – and because, unlike many other versions of this view, Block is steadfast in holding that any theory of meaning must be, in his words, “psychologically relevant”. It must not only explain how representation is possible, but it must also render potential explanations of the causal role and efficacy of states with representational properties. He writes, “a psychologically relevant theory of meaning ought to illuminate the connections between knowing/understanding/learning and usage, on the one hand, and meaning on the other” (BLOCK, 1986, 618).

It is in *An Advertisement for a Semantics for Psychology* that Block looks to develop his psychologically relevant theory. He argues that there is reason to believe that it is only a conceptual role semantics that could satisfy the competing demands of a theory of meaning and a theory of instantiation. Towards this end, Block develops and defends a version of conceptual role semantics that is committed to a number of specific claims about the nature of mental representation and its role in human cognition. Block is, however, careful to note that his account should not be viewed as a theory in itself, but rather as “more of a framework for a theory...” (BLOCK, 1986, 620). Nevertheless, despite the schematic nature of this account, Block develops and defends several specific claims about both the nature of a conceptual role theory of meaning and its instantiation. Amongst the most central of these involve Block’s commitment to a theory of concept acquisition that he suggests is compatible with the demands of both his semantic theory and his theory of instantiation. In this paper, I will argue that Block’s theory fails to live up to its advertised virtues, that it does, in fact, run afoul of the twin constraints that theories of concept acquisition face. How concepts are acquired, on Block’s view, is not compatible with his picture of the nature of concepts and concept possession. Even so, I will maintain that Block’s theory might be revived with little change to the spirit of the original.

In what follows, I will take a closer look at both the demands of conceptual role semantics and the various approaches to acquisition with which it has been thought compatible. In the final section, I will examine Block’s account of how concepts are acquired within his own theory. The remainder of the paper is taken up with an analysis of these views.
A. Conceptual Role Semantics, the short course

Conceptual role semantics was developed first and perhaps most plausibly as a theory of meaning for logical terms or constants. The basic idea is that the functions described by logical operators could be defined strictly in terms of the inferences in which they participated. The meaning of a logical term was thus a function of the set of valid inferences that could be drawn from sentences in which it appeared. Take, for example, a logical term #. How we read the sentence, “The grass is green # the sky is blue,” will depend on what role we take the connective to have. Suppose for instance that we take it to be either the biconditional or conjunction. Determining the adequacy of either reading appears to depend on what inferences we are allowed to draw from sentences in which the connective figures. If it is permissible, for example, to infer that “The sky is blue” from the previous sentence then we can suppose, with confidence, that # does not represent disjunction or the biconditional – and we can get closer to the answer if we know what we may also infer that “the grass is green”.

We would, however, be mistaken if we supposed that these inferences function as a sort of evidence for claims about the logical role of the term tokened in these sentences. What is central to a conceptual role approach to logical terms is that what a given sign or term means is not merely evidenced by its role in inferences, but rather it is constituted by this role. The meaning of a given connective is given by its role in determining the truth conditions of the inferences in which it participates.

The advantages of such an approach to logical terms are substantial. On the one hand the theory offers an account of logical terms that presupposes no representational primitives beyond those that comprise the nonlogical vocabulary of the conceptual system in which it participates.\textsuperscript{2} As a consequence, inferential role theories offer welcome relief to those attempting to find an appropriate reduction base for logical terms and concepts. Moreover, even if we can’t reduce logical concepts to a more fundamental nonlogical vocabulary, we can nevertheless provide an exhaustive overview of what determines the meaning of any logical

\textsuperscript{2} This does not mean, however, that there are no primitives within the theory. What is required are nonconceptual primitives which govern the transitions in the inferences that are made. Our inferential abilities are built out of simple, nonrepresentational, reasoning skills which underlie our capacity to make the appropriate inferences. Peacocke has called these basic inferences primitively compelling, because it is central to this account that the individual find these transitions compelling, but also not answerable to any other inferential process, nor subject to any other proof procedure. See Peacocke (1992, 6). See also Ned Block (1986, 641) for a similar approach. In procedural semantics, these operations are often called “primitive instructions”. See John Haugeland (1989, 66).
connective. Since every inference in which a term participates is itself partially constitutive of its functional role, there can be no applications of the term, within the language, that go against its putative meaning.

What is perhaps most appealing about this approach, however, is the short work it makes of definitions. For while logical terms faced difficulties within a reduction base of empirical concepts, the overarching problem with classical theories of concepts was the inability to provide the requisite conceptual analyses. Since on conceptual role theories terms are not reducible to other more basic terms, there is no requirement that conceptual analysis reveal an underlying conceptual structure for individual concepts. Hence, the theory requires neither the definitions, nor the primitives in which to couch them, that proponents of the classical theory of concepts found so difficult to provide. As a consequence, many philosophers have supposed that a more broadly construed conceptual role semantics might effectively contend with the problems that have confronted analyses of other kinds of concepts. Such an approach would offer an alternative to a reductionist empiricist theory as well as a radical nativism about primitive concepts.³ At its limit, this approach has been adopted by many as a general theory of conceptual content (though often as part of what has been called a “dual aspect” semantics, see McGinn, 1982).

How would such a theory work? There are many different views on how to approach this issue. Most argue, however, that it wouldn’t be much different than the theory given for logical terms and concepts. Conceptual role would be defined in terms of a concept’s inferential or causal role within the larger economy of nonlogical terms. On some views, the conceptual role of a term or concept is specified in terms of the dispositional inferential role of the concept. In other words, the meaning of a term is given by the inferences that one is apt to make with respect to sentences or thoughts that include that term. On most views, the relevant dispositional relations between sentences are not governed strictly by deductive inferences, as they are for logical terms. Instead, these theories tend to include inductive inferential relationships in demarcating the content of a given term or concept, whereas still others cleave to the view that any causal relationship between tokened sentences or thoughts are relevant to fixing the meaning of the terms which comprise them. The move towards a more inclusive inferential or causal apparatus is

---
³ I am moving pretty freely here between a conceptual role semantics for terms in natural language and a conceptual role theory for concepts. I think nothing is lost in this translation. In particular, a CRS for natural language terms specifies that the meaning of the terms is exhausted by the inferences an individual makes or might make involving a given term. The same is true for a CRS for concepts.
necessitated by the view that deductive inferences alone do not cut finely enough to grant unique inferential roles to concepts we would otherwise regard as distinct. Tautalogies are true in every possible world, but clearly they do not mean the same thing in those worlds. As a consequence, a more nuanced and inferentially comprehensive approach has been considered necessary.

Since it is only sentences or propositions that participate in inferences, the conceptual role of a given term or concept would be defined in terms of its contribution to the conceptual role of sentences or thoughts in which it figures. Usually, conceptual role is defined counterfactually – given not strictly in terms of the inferences in which a given concept participates so much as that role it would play if it were tokened. The inferential role of a concept can thus be specified not only in terms of the inferences that an agent actually makes, but also in terms of inferences that would be prompted under circumstances that have not occurred or do not currently obtain.\(^4\)

Perhaps the central difference between a conceptual role theories of logical terms and concepts and the broader account of concepts sought after here is the relative difference in our ability to specify the component inferences that give the meaning of these terms. Since the logical terms form a closed set of interdefinable functions, it is fairly easy to demarcate what inferential properties are indicative of which functions. One can define, for example, the material conditional in terms of negation and conjunction, and conjunction in terms of negation and the material conditional, etc. Similar definitions are available for the quantifiers as well as modal operators. In these cases, then, one can straightforwardly isolate the conceptual role of the terms and concepts involved by merely adjudicating the inferences in which the concepts participate.\(^5\) The same is not true for our broader theory. Nonlogical terms and concepts participate in a seemingly endless array of actual or possible inferences that appear relevant to determining their conceptual role. Take, for instance, the concept UNLOVEABLE.\(^6\) In many instances, this concept may be associated with inferences concerning one’s looks, financial status, religion, ethnicity, sense of humour, hair and/or eye

\(^4\) It remains unclear whether conceptual role should be understood in terms of idealized reasoning capacities or rather in terms of what people actually do. I won’t distinguish between these views for the purpose of this discussion, though the former may prove difficult to square with our naturalism.

\(^5\) The point is not that we can offer for any particular logical term a succinct definition of its conceptual role, but rather that we can specify for each logical term a finite set of inferences which exhaust the transitions which that term licenses.

\(^6\) I will follow the convention that references to concepts be placed in capitals.
colour, sleep patterns, tastes in music, art, movies, shoe size, etc. It is an all too common practice to infer from, say, X is a graduate student, to X is unloveable. If we are to count each of the inferential associations of a given concept as a determinant of the meaning of that concept then the task of isolating the inferential role of these concepts, of determining what constitutes their meaning, appears impossibly complex.\footnote{For many philosophers, the specification problem is at the heart of their rejection of conceptual role semantics. See, for example, Stephen Schiffer’s discussion in Remnants of Meaning (1987, 109). It strikes me, however, that our inability to solve the specification problem is not clearly relevant to the acceptability of conceptual role theories in general. For while it might prove difficult to disentangle, for epistemic reasons, the various strands of a concept’s inferential relationships, our difficulty in determining these relationship has no bearing on whether or not they are constitutive of meaning.}

The complexity of this problem does not, however, imply that it is in principle impossible to specify the inferential associations, or liaisons, of a given concept. However, even if we can supply the requisite analyses of inferential properties of a given conceptual role, it remains unclear as to which of these properties are essential to the possession and acquisition of the concept itself. Consider again the concept UNLOVEABLE. The inferences I am prepared to make from statements involving this concept may differ greatly from yours, so much so that we might believe that our concepts of love are almost unrecognizably distinct. The same may hold for my future selves. After all, our tastes change as we age – and thus not only what is lovable to me at one time might seem banal and insipid at another, but, it appears, what I mean by lovable, or unloveable will also change. Insofar as we allow alteration and expansion within our inferential equations, we seem to be adopting different concepts. If there are aspects of the inferential role of your concept UNLOVEABLE that differ from the inferential role of my concept, does that mean we have different concepts?

Some conceptual role semanticists have argued that it doesn’t show that there are different concepts at work, merely different inferential liaisons attached to the same concept. Christopher Peacocke has championed a version of this view. He argues that there is a criterion, the Fregean criterion of informativeness, which allows us to distinguish between meaning-constituting inferences and those that are not relevant to the semantics of the term. Thus, it might be the case that red hair renders someone unloveable, on a particular perspective, but it is not essential to the meaning of the term. The reason, for Peacocke, is that the equation of red hair and unloveability is informative. It is cognitively significant (1992, 2). The equation, however, of unloveability with the incapacity for being loved is redundant. They are intersubstitutable.
without any change in the informativeness of the sentences in which they appear, and thus, on Peacocke’s theory, they are the same concept. What is central to views like Peacocke’s is that there is some criterion by which we can delimit those inferences that are criterial for the possession of the concept in question. As a consequence two individuals can have the same concept even though the conceptual roles of their concepts are not the same. What must be true, however, is that both individuals are inclined towards the same criterial inferences.

Ned Block disagrees. He argues that the absence of an analytic/synthetic distinction forces us to do away with what he calls “the crude dichotomy of same/different meaning in favour of a multidimensional gradient of similarity of meaning” (BLOCK, 1986, 629). Block takes Quine’s refutation of the analytic/synthetic distinction to imply that there are no purely criterial inferences. As a consequence, he cannot suppose that there is some select subset of a concept’s inferential properties that would be prerequisite for the mastery of a given concept. Instead, Block assumes that each of a term’s inferential properties is at least partially constitutive of the meaning of that term. The meaning of any concept is thus a function of its entire conceptual role.

The price of this ecumenism is severe, however. Conceptual roles are typically defined relative to other conceptual roles within a representational system. This is easy to see. We individuate the inferential role of a given sentence or thought in terms of the sentences or thoughts to which it is inferentially related. These sentences or thoughts are likewise inferentially related to others within the system. The problem is that what inferences one is inclined to make depends on what other sentences or thoughts are instantiated within the system itself. The belief that poverty is a sin prompts the belief that current economic policies are evil only as a consequence of some other belief that suggests that these policies promote poverty. Clearly, not everyone shares this belief, or indeed a multitude of others that are relevant to the inferential roles of these sentences and their component parts. In supposing then that all of the inferential properties of a given concept or sentence are relevant to the meaning of that concept or sentence leads one to embrace not only holism, the idea that the meaning of a given sentence or thought depends on the meanings of all other sentences and thoughts within the system, but also the idea that content is idiosyncratic. That is, since no two people share the same beliefs, no two people share the same concepts.

To put it another way, one can reasonably deny that unloveability and red hair go hand in hand, but one cannot sensibly deny that the incapacity for being loved does not go with unloveability. A denial of this latter claim is to give up on the concept altogether.
This result is potentially devastating for any theory of concepts that is motivated by the hope of explaining common sense psychological generalizations. Since no two individuals are likely to share the same concepts, it is unlikely that their behaviour can be subsumed under generalizations that advert to the contents of their beliefs or desires.

It is Block’s hope that insofar as these problems arise, they might be defused by an appeal to some notion of semantic similarity. After all, it must be noted that the vast majority of an individual’s beliefs are not so different than one’s neighbours. That is, most people will share basic sensory beliefs, and basic biological desires – not to mention, in some large degree, a basic agreement concerning the laws of inference that are to be applied to beliefs and desires. Any difference in inferential role must then be defined against this background of a shared intellectual outlook, partly biological and sensory, and partly communal. It is perhaps possible to define some metric of similarity that governs how we attribute beliefs and desires that are appropriately similar to each other. These laws would not be, of course, exceptionless, but of course psychological laws aren’t exceptionless either.

However, even if we suppose that some similarity index might be employed to form the backdrop against which useful and interesting psychological generalizations can be made, Block’s holism encounters difficulties on a wholly different front. The attempt to develop a theory of concept acquisition which at once respects his holism as well as his picture of evidence and learning tests his similarity metric in some rather serious ways. In what follows, I will first delineate Block’s account of concept acquisition before turning to a discussion of the ramifications of his holism for this view.

B. Cognitivism

A cognitivist theory of concept acquisition is committed to the idea that the majority of one’s concepts are acquired through inductive belief formation processes. On these views, coming to acquire a concept is entering into the appropriate belief state with respect to the semantic properties of a given concept. On some views, this has been defined in terms of learning specific definitions or necessary and sufficient conditions for the application of a given term or concept, but there is considerable latitude in the details of these theories to allow for significant disagreement. What all such cognitivist theories require, however, is simply that concepts are learned and that coming to possess a concept involves coming to have particular beliefs.

Paradigmatically, how we learn is a function of forming and confirming the appropriate hypotheses (see Fodor, 1981b). Typically, an individual
develops a hypothesis about the meaning (or, in this case, the conceptual role) of a given term or concept and tests that hypothesis against the relevant features of their future experience. What makes such learning processes rational is that the relation between the hypothesis confirmed and the concept acquired is fundamentally an evidentiary one. The confirmation of the hypothesis justifies the subsequent use or application of the concept.

While there has been very little attention paid to the issue of concept acquisition within the conceptual role semantics literature, it has been widely assumed that the appropriate account of concept acquisition for these theories would be noncognitive. Indeed, proponents of concept role theories have been very careful to disassociate themselves from the standard learning models of concept acquisition. The reason for this has been twofold. In the first place, since the acquisition of a given concept amounts to nothing more than acquiring a certain functional or inferential organization, it has been thought that there is no need to take on any additional explanatory commitments involving the possible avenues by which this organization is achieved. What matters is thus not how we come to acquire a certain inferential organization, but rather whether such an organization adequately describes the relevant semantic relationships. The second concern stems from the identification of conceptual role theories with what have been called use theories of meaning. On these views, to acquire a concept is not to come to learn a definition or what it stands for, but rather it is the acquisition of a sort of inferential capacity or ability. The distinction between knowing how and knowing that has often been invoked to underscore this distinction (a distinction that has now met with some serious skepticism, see STANLEY and WILLIAMSON, 2001). Acquiring an ability, like riding a bike, does not require the mastery of any set of propositions, but rather the acquisition of a particular set of abilities or capacities. On this model, a concept is perceived as a sort of skill and it is, as a consequence, no more definable or cognitively accessible than the skills of great home run hitters (see, for example, LEWIS, 1990, 519; DEVITT, 1996, 52).

The move to a noncognitive, nonrational approach to concept acquisition provides the conceptual role theorist with a convenient solution to the problems raised by what Fodor has called “the standard argument” (FODOR, 1981b). The standard argument suggests that any rational learning model of concept acquisition is committed to the view that at least some, and in Fodor’s view, quite possibly most, concepts are innate. The argument turns on the fact that the formation of any hypothesis will require, on some level, a specification of its contents in terms of concepts which are already available to the
subject. These concepts, at least initially, cannot be learned in the usual way.

However, noncognitivism offers a putative means to avoid the most damaging of Fodor’s conclusions by giving an account of concept possession and acquisition that does not depend on rational learning strategies. On this view, since concepts are acquired through feedback mechanisms which constitute neither evidence nor justification for how that concept is used, there is no presumption that acquisition proceeds through hypothesis testing procedures and hence no demands for a proprietary language in which to couch those hypotheses. If we assume, as many conceptual role semanticists do, that acquiring a concept is akin to acquiring a skill or a knack, then one can offer a genitive theory of concepts that does not presuppose any innate conceptual repertoire. That is, since acquiring a concept is a form of learning how, rather than learning that, it is no longer apparent that such views fall prey to the standard argument’s conclusion.

This argument is, however, only as satisfying as our story of noncognitivist concept acquisition, and it is not clear that we have one of those. For even if the demands of hypothesis testing and confirmation no longer apply, we nevertheless lack any robust sense of what mechanisms a noncognitivist theory of concept acquisition does require. Fodor has argued that while we might plausibly maintain that learning-how does not require hypothesis testing, there remain some good reasons to suppose that it might after all (FODOR, 1998, 124-125). We are, if anything, less prepared to offer a fully-fledged theory of noncognitive concept acquisition than we are for the impoverished cognitive side. If the noncognitivist believes that learning-how might afford us a sort of back door out of the standard argument, it is not obvious that he has a key.

And Block wants to go through the front door. Block’s holism is unusual in supposing that a cognitivist learning theory is required by conceptual role semantics, but in light of the concerns raised above, a direct, rational, approach to concept acquisition offers some distinct advantages. However, even granting this, we must acknowledge that such a theory cannot be complete, for at some level it presupposes that there are some concepts available for forming the requisite hypotheses. The genesis of these concepts cannot be explained in recourse to rational learning strategies. As a consequence, the approach Block advocates here cannot be the sole story of concept acquisition for a conceptual role theorist. Some concepts must be acquired outside of rational learning strategies, either in the manner of skills or physical capacities (provided this can be worked out) or as a consequence of the development of innate conceptual roles.
Despite this, there are some advantages to adopting even a partially
cognitivist theory of concept acquisition. For what does seem true is that
what concepts an individual possesses is a function of their educational
and experimental background. Insofar as we acquire evidence for our
beliefs, we also, it seems, increase our capacity to consider them. We
might say growth in knowledge correlates very well with the growth in
the breadth and complexity of one’s conceptual repertoire. For example,
the concepts that are available to a physicist in his or her description of
the universe seem to arise as a consequence of his or her directed study
into those specific domains of discourse. The same is no less true in other
disciplines, inside and out of the sciences. The concept, for example, of
a FAKEY comes most easily to those who are versed in skateboarding,
while the concept F-STOP is usually mastered only by those who possess
the appropriate collateral knowledge within photography. Indeed,
possession of the appropriate collateral knowledge is often taken to be
criterial for ascription of concepts of specific sorts. The physicist does
not, so it seems, acquire the knack for thinking about quarks, so much as
the requisite propositional knowledge. What noncognitivist theories often
lack is any way of explaining the relationship between what an agent
knows and the concepts he or she possesses. As a consequence, Block
might rightly assume that a conceptual role theory that is compatible
with a cognitivist account of concept acquisition gains a measure of
empirical plausibility. The opposite also seems true. Theories of concept
acquisition which cannot account for the interrelationship between what
an individual knows and what concepts they possess become more
questionable on precisely these grounds.9

The substance of Block’s theory, developed largely in reaction to
Fodor's version of the standard argument, is that concepts are acquired not
through the formation of speculative definitions, subsequently confirmed,
but rather through the formation and confirmation of hypotheses
concerning the functional role of specific terms or concepts.10 He writes,

---

9 The same advantages accrue to cognitivist theories of concept possession. For on this
view, the conceptual role of a term is determined by the set of inferences in which it
participates. However, the inferences to which a concept or term makes a contribution
will depend, at least in part, on the collateral knowledge that the subject possesses.
As Fodor points out, “you can’t identify a dog by its barking unless you know(/believe)
that dogs bark.” And, “surely, you won’t infer from dog to animal unless you know(/believe)
that dogs are animals” (1998, 125).

10 Block does not address the question of how it is that hypotheses about conceptual roles
are confirmed, but one could imagine that such an hypothesis is confirmed, in the case of
natural language terms, by comparing one’s usage of that term against its employment
by others. The story is apt to be different for concepts that are not acquired concurrently
with equivalent expressions in natural language. One could imagine, however, that these
hypotheses are tested against instances of the properties they are meant to represent.
“to the extent that hypotheses are involved, they are hypotheses about how a term functions in thought, reasoning, problem solving and so forth” (BLOCK, 1986, 647). Since these hypotheses are meant to capture the conceptual role of a given term or concept, they appear to avoid the consequences of Fodor’s standard argument. Since the conceptual role of a term catalogs the inferential relationships in which a term participates, we might suppose that it offers something very much like a definition of the term in question. However, Block maintains that this view is misguided. Insofar as conceptual roles chart the function of a term or concept in thought, they do not straightforwardly correspond to standard conceptions of definitions.

Block makes his case for this model through the use of an example. He notes that in acquiring the concepts of elementary physics, he did not learn definitions of the terms, but rather he learned how to use the new terminology. “I learned certain relations among the new terms themselves (e.g., the relation between force and mass, neither of which can be defined in old terms) some relations between old terms and new terms, and, most importantly, how to generate the right numbers in answers to questions posed in the new terminology” (BLOCK, 1986, 648). Block invokes Lewis’s functional definitions as a model of what is acquired in this learning process (see LEWIS, 1980). We do not learn a static definition of what a term means, an analysis into other more fundamental terms, but rather an explication of the role of a term within a theory.

If Block avoids the charge that he is in actuality presupposing that terms can be given the sorts of definitions Fodor finds objectionable (and it is not clear that he does), it nevertheless remains clear that the account he is offering is cognitive in character. Block supposes that we can come to acquire new concepts by learning the role of a given term within a larger theory, where the mechanisms for picking out this role are themselves cognitive. We form hypotheses about the role of, say, QUARKS within the larger theory of quantum mechanics. Whether we perceive these hypotheses as definitions in the classical sense is debatable, but it is clear that in forming hypotheses we are employing concepts that we antecedently possess. These hypotheses are subsequently confirmed or disconfirmed and through this feedback mechanism we alter our hypotheses accordingly.

What makes Block’s view unusual, as a conceptual role semanticist, is his acceptance of Fodor’s contention that hypothesis formation and confirmation currently constitute the only plausible model of the mechanism underlying cognitive theories of concept learning and acquisition (BLOCK, 1986, 647). However, he does deny the attendant claim that in acquiring concepts we must form definitions in terms of
concepts we already possess. There are two reasons for this denial. As we have already seen, the hypotheses confirmed are not traditional definitions. In forming hypotheses we speculate on the role of a term within a given theory, not on its dictionary definition. More importantly, perhaps, Block suggests that in acquiring new concepts or terms we often employ terms with which we are only marginally familiar. It is often the case that the fundamental causal or inferential relationships relevant to acquiring a new term involve the concurrent acquisition of other terms. One cannot acquire the concept ‘momentum’ without the concept of ‘acceleration’, nor perhaps ‘bachelor’ without the concept of ‘marriage’. Some terms are learned holistically, in conjunction with a family of largely inferentially interdefinable terms. However, even if this is often the case, it is nevertheless apparent that it is central to distinguishing and demarcating the functional roles of these terms to frame them, and the relationships that exist between them, in terms that are already understood.

C. Can there be a holistic cognitivism?

Block’s theory is thus at once holistic and cognitive in character. He holds that the inferential relations (or liaisons, as they are sometimes called) of a given concept cannot be divided into essence and accident, and as a consequence each of the inferences in which a concept participates is partially constitutive of its meaning. However, since the inferential liaisons of a given concept are inextricably tied to the inferential relationships of a multitude of other concepts, and these with a multitude of others, it appears that the meaning of one concept is likely to be a function of a good many if not all of the others. This is the sum of Block’s holism.

What his cognitivism requires, however, is more complex. For his supposition that new concepts are acquired through rational learning strategies presupposes a confirmation relation between a given hypothesis and the concept whose functional role the hypothesis describes. A confirmation relation is, at bottom, an epistemic relation between two sentences or propositions. As a consequence, the development of any hypothesis will require that the agent employ, at some level of description, some (and quite possibly all) of the concepts that he has previously acquired.

If the combination of these two doctrines, holism and cognitivism, helps deflect criticisms of conceptual role semantics by tying the rational acquisition of concepts to increases in general levels of knowledge, it nevertheless introduces new tensions that are unique to this approach. These emerge when we consider the nature of the hypotheses that are
employed in acquiring new concepts. Notice that Block is committed to the view that insofar as each hypothesis employs concepts that the individual already possesses, it potentially employs every concept that the individual has. This is a consequence of his holism. If he accepts the general claim that each of a concept’s inferential liaisons is partially constitutive of the meaning of a given term or concept, then it becomes clear that any hypothesis will involve all of the concepts which participate in those liaisons. In the limiting case, this is every concept that the individual possesses (see FODOR & LEPORE, 1992, for a version of this argument). 11

The problem for Block is simply this: if the meaning of every concept is a function of every other concept within a given conceptual system, then the acquisition of a new concept will change the meaning of the concepts in the confirmed hypothesis. The meaning of the concepts in the hypothesis are a function of the inferential liaisons they shared with each of the other concepts that the individual possessed prior to acquiring the new concept. Once a new concept passes through the confirmation regimen 12, each of the old terms acquires new inferential liaisons – that is, new inferential roles with respect to the concept acquired. This changes the meaning of each of the terms that comprised the hypothesis – and each of the terms to which they were inferentially related. On any holistic view, of course, this is all of them.

What this suggests is that the hypothesis which was confirmed in the acquisition of the new concept is literally inexpressible in the conceptual system in which the new concept exists. The reason should be clear. Block’s holism requires that the inferential role of any concept be a function of its inferential liaisons with each of the concepts within the system. However, with the addition of a new concept or concepts, the inferential liaisons of each of the pre-existing concepts must expand to include their inferential relations with the new concept and the changes in inferential relationships that stem from its inclusion within the system. As a consequence, the inferential role that the hypothesis embodied no longer exists within the system itself.

11 Some have argued for what has been called a molecular conceptual role holism, which suggests that concepts divide into relatively broad categories in which inferential roles are co-determinative of the meaning of the concepts within those categories. Logical terms and concepts might, as we have seen, constitute one plausible candidate for a molecular or indeed might the concepts of physics and chemistry. Certainly, within these categories, there is an acceptance of a limited holism, but shy of a complete theory of conceptual role we cannot know the substance of the claim that these inferentially isolated islands of meaning might avoid the objections I raise here.

12 This confirmation regimen will depend on a number of factors – the complexity of the conceptual role in question, the conditions under which it is confirmed and their frequency, etc.
Moreover, the new concept cannot have a functional role within the new system which is the same as that in the confirmed hypothesis since the conceptual role that the hypothesis specified cannot exist in a conceptual system with the concept that has been acquired. The problem is that if we accept a hypothesis formation and confirmation account of concept acquisition, we must also accept that the functional role of the new concept is the same as the hypothesis. For the confirmation relation is itself an inferential relation, one which licenses inferences from the conceptual role of the hypothesis to the conceptual role of the new concept. However, the conceptual role of the hypothesis can make no allowances for the inclusion of the new concept without a concurrent change in its meaning.

What this suggests is that Block cannot simultaneously maintain a cognitivist view of concept acquisition while holding onto his holism. For what cognitivism demands is that there be some rational relationship between the content of the concept that is required, given by its inferential role, and the confirmed hypothesis. In particular, the confirmation relation that exists between the hypothesis and the concept acquired should allow one to infer from the hypothesis to the concept, for what is confirmed is that the hypothesis adequately or correctly characterizes the new concept. But there is no guarantee that this is the case as the new concept does not have the inferential role of the confirmed hypothesis. This breaks the evidentiary connection between the confirmation regimen and the subsequent application of the concept. As a consequence, we cannot suppose that the concept acquired retains the same content as the hypothesis.

D. Can there be a holistic noncognitivism?

In the previous section, I argued that Block’s holism demands that the content of each of an individual’s concepts change in response to the acquisition of a new concept, while his cognitivism requires that the new concept retain the conceptual role of the hypothesis through which it was acquired. Some have argued, however, that the problems similar to those encountered above apply quite generally to any theory of concept acquisition for a holistic theory of meaning. As we have seen, the acquisition of any new concept will change the meaning of every other concept within the representational system. What this suggests is that there can be no incremental means of acquiring a concept within a holistic theory of meaning. Michael Dummett has made this point with respect to learning terms within a natural language.
The situation is essentially similar to that of a language all of whose sentences consist of single words, i.e. have no internal semantic structure;...it becomes unintelligible how the speakers of the language could ever have come to associate...senses with their unitary sentences... In the same way, if a total theory is represented as indecomposable into significant parts, then we cannot derive its significance from its internal structure, since it has none... (DUMMETT, 1973, 599-560).

The problem is that a holistic language must, in some sense, be acquired all at once. Any incremental procedure would demand that some concepts or terms be acquired before the others. However, it is not clear how one can acquire a subset of a given conceptual system without acquiring the rest of that system at the same time. That is, if the meaning of any concept in a system depends on its relationships, however defined, with each of the other concepts in that representational system, then it appears impossible to acquire any of the concepts within that system without simultaneously acquiring all of them.

Perhaps this explains Block’s emphasis on acquiring sets of concepts concurrently. “In my own case, I heard a large number of unfamiliar terms more or less all at once: ‘mass’, ‘force’, ‘energy’, and the like. I never was told definitions of these terms in terms I already knew” (BLOCK, 1995, 402). Whether Block’s case reflects our own experience doesn’t matter. The overriding point here is that in acquiring a set of new concepts Block does not appeal to concepts that he has previously possessed. Instead, he argues that the acquisition of these concepts was achieved by systematically relating them to each other.

It is unlikely however that a genuinely holistic theory of meaning could rely on such devices. How the various concepts in Block’s Newtonian equations are related is a function of concepts Block already possesses (That is to say, the hypotheses Block forms about how ‘mass’ and ‘force’ are related will depend on his ability to use concepts that capture that relation). What Block supposes, that a holist cannot, is that the meaning of the concepts within a given domain (in this case, Newtonian mechanics) are exhaustively determined by the inferential relationships between the concepts in that domain (He must also suppose that the relationships that he describes between these concepts also form a closed circle of interdefined terms or concepts).

We might think of it this way. There are two sets of concepts we need to consider here. There are those of Newtonian mechanics, which are defined through various operations that Block learns to perform over them. If he learns how to perform these operations through hypothesis testing and confirmation, then he is also committed to the existence of a set of concepts which express the nature of these relations. If these
latter concepts formed a closed domain, then it appears possible that they could be used to acquire concepts in another closed domain, Newtonian physics. These approaches to concept acquisition appear to eschew holism in favour of a more molecular approach to the semantics of concepts. Despite Block’s apparent, though halting, endorsement of holistic theories, his examples suggest that he is at least sometimes inclined towards a molecular conceptual role holism, which suggests that concepts divide into relatively broad categories in which inferential roles are co-determinative of the meaning of the concepts within those categories. Logical terms and concepts might, as we have seen, constitute one plausible candidate for a molecular or indeed might the concepts of physics and chemistry. Certainly, within these categories, there is an acceptance of a limited holism and the same issues recur on the smaller scales that the theory advances.

The smaller scale also fails the larger one. For while a molecularism stills face problems involving the relations between each of the terms within the domain, it also amounts to a denial of an overarching holism. What it suggests is that within a given representational system, there are conceptual domains which are independent of the other concepts within the representational system.

E. Whither Similarity?

The above argument is apt to seem excessively uncharitable to proponents of holistic theories of meaning. For what this argument appears to neglect is that on any incremental account of concept acquisition, whether it be cognitivist or otherwise, new concepts will effect only relatively minor changes in the meaning of the other concepts or terms within the system. On a conceptual role semantics, what this means is that while the expansion of the conceptual system will result in changes in the meaning of terms with which the new concepts are inferentially related, it will not effect wholesale revisions in the inferential relationships between the concepts that already exist within the system. What this suggests is that the meaning of the terms or concepts within the previous system will not be significantly altered. As a consequence, one might argue, conceptual role semantics can live with holism on the grounds that the meaning of concepts within the systems will remain roughly commensurate will the concepts in the system prior to the acquisition of a new concept or concepts.

This view is reinforced by the relatively banal observation that the preponderance of our beliefs are grounded in simple experiential or observational terms. These concepts, which are grounded in biologically basic brute experiences form much of the basis of our conceptual palette.
and they embed. These are likely to be universal, or largely so, and hence relatively immune to wholesale changes in our conceptual scheme.

We find in this a second opportunity to deploy Block’s notion of meaning similarity in defense of conceptual role semantics. For what Block has suggested previously is that we must do away with the idea that particular concepts are to be associated with an all or nothing criterion for concept possession or understanding, and instead opt for what he calls a “multidimensional gradient of similarity of meaning” (BLOCK, 1998, 629). We might, then, be disposed to ascribe the same concept to individuals with conceptual or inferential roles that are not identical. Moreover, we could expect the same level of charity within an individual, and across concept acquisition. Thus, we needn’t expect that the acquisition of a new concept will, fundamentally, change the semantic resources available to the individual, for if we suppose that the system has achieved a relative degree of inferential depth and complexity13, then insofar as the acquisition of new concepts will result in changes within the inferential structure of a conceptual role semantics, these changes are likely to be localized largely within the conceptual domain of the new concept or hence inconsequential within the framework of the entire system.

Each of these proposals is tied to the idea that concept identity depends requires that both concepts share exactly the same inferential roles. However, public meaning may require something substantially less. Ingo Brigandt has argued that concept identity is defined not strictly in terms of their inferential roles (which may be more or less the same), but rather by how those roles satisfy our semantic interests. Hence, two concepts may be reckoned identical even if they do not share precisely the same inferential liaisons. How a concept is used determines the criterial basis for the successful application of a given concept relative to our interests. Brigandt writes, “these interests determine how a concept is to be individuated; and as the same term can be subject to different

---

13 The point of this rider is to note that smaller conceptual systems will be subject to far more instability with the addition of new concepts. Children are much more likely to experience a paradigm shift in their conceptual growth than adults, if only because their conceptual system is subject to more frequent and inferentially onerous changes than adults. (Consider for example the child’s first steps towards making an inferential distinction between concepts of having and possessing and the relatively complex concepts of ownership and purchasing. Economic concepts postdate merely possessive ones, but in their acquisition they change the child’s conceptual relationship with virtually every object in their environment.) Nevertheless, adults possess no special immunity to large scale shifts in the inferential organization of their conceptual systems and as a consequence cannot suppose that newly acquired concepts might not threaten the intertranslatability of one’s current and past representational systems in serious ways.
philosophical studies and theoretical interests, its content may be individuated in different ways" (BRIGANDT, 2004, 32). Hence, a slippage in concept meaning can be at least partially regulated by how they serve our semantic interests with their public expression.

The problem remains, however, that in the absence of any specific proposals concerning the nature of conceptual roles, it is impossible to quantify the degree to which any two conceptual roles are similar. Simply put, if we can’t say precisely what the conceptual role of a given concept amounts to, then we are even harder pressed to say what it would take to have a conceptual role with a similar content. Moreover, even under such a theory, it might well remain unclear how similar two conceptual roles would have to be to warrant the ascription of the same concept or what contexts would be relevant to determining whether such ascriptions would be justified. The promise of conceptual role semantics will depend in large part on how effectively these issues can be resolved.

These worries are not insubstantial, but despite this, the idea that two distinct conceptual roles may nevertheless possess roughly equivalent conceptual content suggests that Block’s theory of concept acquisition might be resurrected. Insofar as new concepts change the contents of other concepts within the system, such changes would not, in general, significantly change their meaning. However, since the acquisition of any new concept will still change, to whatever degree, the inferential roles of the concepts that compose the confirmed hypothesis, it is no less true that the evidentiary connection between hypothesis and the conceptual role acquired is broken. The concept acquired is not the sum of its hypothesized parts. As a consequence, even on a hypothesis testing model of concept acquisition, we must accept that concept learning is not a strictly rational process.

Despite this, insofar as this process reflects relative similarities between the conceptual roles that are acquired and the concepts that comprise the relevant hypotheses, it will remain true that concept acquisition will appear largely rational. We might consider the learning process here pseudo-rational. For new concepts will be acquired through what appear to be rational learning strategies. Hypotheses are formed and tested against experience, but the acquisition of the concept itself changes, in most cases we might suppose quite slightly, the conceptual makeup of the entire system. Indeed, if it were a regular feature of new concepts that they came radically apart from the confirmed hypotheses that led to their acquisition, we would be unlikely to pursue the same strategies for other concepts. Insofar a conceptual role determines meaning, it faces simultaneous referential demands that rational learning strategies effectively satisfy. Thus, no matter in how many directions the
new thread in the rug is pulled, its acquisition is tied specifically to the hole it was meant to fill. Our acquisition strategies are apt to reflect this concern. Hence, the “gestalt switches” expected are not the earthquakes that Dummett (or perhaps even KUHN, 1970) would suppose them to be, for while new concepts change the system, we have no reason to suppose that they undermine or alter in any fundamental way the overall inferential structure of the system.

This does not mean, of course, that such changes are in fact impossible. Certainly, children experience conceptual changes that radically reorient their conceptual understanding of the world. Perhaps Einstein’s theories had much the same impact within the relatively closed conceptual world of physicists and chemists. Thus, we must also contend with the possibility that the acquisition of new concepts could in fact radically change the existing inferential relationships within a given conceptual system. However, in the absence of a specific theory about the nature of conceptual roles, it is difficult to imagine the conditions under which new concepts might lead to substantive changes within a holistic representational system, but it seems clear that nothing about Block’s theory rules out such a possibility. Indeed, this is precisely why we cannot suppose there is a rational link between our practice of acquisition and the concepts that are acquired. For there is no guarantee under current conceptions of inferential role semantics that some new concepts will not dramatically change the representational capacities of the system itself.

Moreover, it is easy to imagine circumstances in which such changes could occur. Take, for example, a simple conceptual system of the sort Wittgenstein envisages at the start of the *Philosophical Investigations* (1953). In this language, the builder and his labourers have but four words, “slab”, “beam,” “block” and “pillar”. Each word is in fact a shorthand for a sentence, one of the sort, “bring me a slab,” or “bring me a pillar”. If the meaning of each term, on such a simple system, depends in part on the meaning of every other term in the system, then it appears likely that a new concept, even one conceived in terms of these existing concepts, will radically alter the inferential relationships between the various terms. And so we might find, at least for inferential role semantics, that size does matter. Smaller representational systems would seem more susceptible to a kind of discontinuous semantic development. New terms and concepts will have a greater impact on the inferential associations or relationships in smaller systems than in larger ones.

We can see evidence of this in the semantic development of children. They are far more likely to experience a paradigm shift in their conceptual growth than adults, if only because their conceptual system is more circumscribed and thus more subject to frequent and inferentially
onerous changes than adults. Nevertheless, this does not imply that adults are immune to large scale shifts in the inferential roles of their representational systems. However, what factors are relevant to these sorts of changes in adults are less obvious.

Imagine a holistic theory of flavouring. On this view, the taste of a given dish is a function of each of the flavours from which it is comprised. The addition of new flavours will not alter the existing pattern of tastes (or rather the taste sensation these ingredients produce) within a dish even though it will change how the entire dish tastes. In the same way, the taste of something is not merely the sum of its parts, but rather the sum of these parts as a consequence of their relations to each other. We might be able to imagine, from consideration of the ingredients alone, what something might taste like, but we are often surprised, because we find it difficult to imagine particular flavours in combination. Consider for example the recent vogue of chocolate made with chili peppers. On paper, this is a poor match, since one is typically savory, while the other sweet. Despite this, these mutant taste combinations are often both surprising and compelling.

Despite this, there are general rules for adding flavours and general suppositions about what something might taste like. These rules of thumb are usually relatively accurate about the nature of flavour combinations, but they needn’t always be so. Sometimes the introduction of a new ingredient might result in a radically new and unexpected flavor. A map or theory of these changes is the appropriate subject matter for food scientists and psychologists, and in the same way a theory about the sorts of inferential changes wrought by the addition of new concepts is the subject matter of a more robust conceptual role semantics. The size of the conceptual system is but only one factor within the inferential stew that Block is trying to assemble.

However, if we can’t predict, at least at present, what sorts of factors are relevant to determining the impact of a new concept or concepts within a conceptual system, then it seems clear that we can’t presuppose that there is a wholly evidentiary relationship between the hypotheses formed and the concepts that are acquired. The relationship between a concept acquired and the hypothesis from which it was confirmed cannot thus be divined in any straightforward manner. And what this suggests is that many of the criticisms that have been levied at Block’s theory have been misguided. In particular, Fodor has argued that rational learning strategies imply that the agent acquires new concepts by learning a definition in the form of the confirmed hypothesis. However, he notes,

14 Sue Carey charts some of these differences in her (1982).
agents who have mastered the use of particular concepts are rarely able to define them in more basic or fundamental terms. As a consequence, Fodor assumes that concept learning cannot be based on rational learning strategies.

What these arguments show, however, is that rational learning strategies can be used in the acquisition of new concepts, but that the addition of the new concept within the conceptual system makes the definition of the concept (the confirmed hypothesis) no longer expressible in the language of the conceptual system. We cannot give precise definitions solely because we no longer possess the means with which to give them. Nevertheless, we are able, usually, to go quite some ways towards expressing the meaning of the concept in other terms. This is precisely what we should expect on the account given above. If the old concepts of the hypothesis evaporate with the acquisition of the new concept, they nevertheless leave a residue with which we can trace out the broad outlines of their previous existence.

References


______. “Holism, Analyticity, and Inferential Role”. MS. (forthcoming).


Recebido em agosto de 2012.

Aceito em dezembro de 2012.